



THE UNIVERSITY *of* EDINBURGH

## Edinburgh Research Explorer

### A Genome-wide Association Analysis of a Broad Psychosis Phenotype Identifies Three Loci for Further Investigation

**Citation for published version:**

Psychiat Genomics Consortium, Psychosis Endophenotypes Int Conso, Wellcome Trust Case-Control Consor, Bramon, E, Pirinen, M, Strange, A, Lin, K, Freeman, C, Bellenguez, C, Su, Z, Band, G, Pearson, R, Vukcevic, D, Langford, C, Deloukas, P, Hunt, S, Gray, E, Dronov, S, Potter, SC, Tashakkori-Ghanbaria, A, Edkins, S, Bumpstead, SJ, Arranz, MJ, Bakker, S, Bender, S, Bruggeman, R, Cahn, W, Chandler, D, Collier, DA, Crespo-Facorro, B, Dazzan, P, de Haan, L, di Forti, M, Dragovic, M, Giegling, I, Hall, J, Iyegbe, C, Jablensky, A, Kahn, RS, Kalaydjieva, L, Kravariti, E, Lawrie, S, Lins-Zen, DH, Mata, I, McDonald, C, McIntosh, A, Myin-Germeys, I, Ophoff, RA, Pariente, CM, Paunio, T, Picchioni, M, Sussmann, J & Whalley, H 2014, 'A Genome-wide Association Analysis of a Broad Psychosis Phenotype Identifies Three Loci for Further Investigation', *Biological Psychiatry*, vol. 75, no. 5, pp. 386-397.  
<https://doi.org/10.1016/j.biopsych.2013.03.033>

**Digital Object Identifier (DOI):**

[10.1016/j.biopsych.2013.03.033](https://doi.org/10.1016/j.biopsych.2013.03.033)

**Link:**

[Link to publication record in Edinburgh Research Explorer](#)

**Published In:**

Biological Psychiatry

**Publisher Rights Statement:**

Copyright © 2014 Elsevier Inc.

This document may be redistributed and reused

**General rights**

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

**Take down policy**

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact [openaccess@ed.ac.uk](mailto:openaccess@ed.ac.uk) providing details, and we will remove access to the work immediately and investigate your claim.



## A Genome-wide Association Analysis of a Broad Psychosis Phenotype Identifies Three Loci for Further Investigation

Psychosis Endophenotypes International Consortium<sup>\*,a,b</sup> and Wellcome Trust Case-Control Consortium 2<sup>a,b</sup>

\*Address correspondence to Elvira Bramon, M.D., Ph.D., Mental Health Sciences Unit & Institute of Cognitive Neuroscience, University College London, London W1W 7EJ, United Kingdom

<sup>a</sup>A full list of individuals who qualify for authorship, along with their affiliations, is provided in the Acknowledgments

<sup>b</sup>A full list of members of the PEIC and WTCCC2 is provided in the Acknowledgments

Received October 23, 2012; Revised March 27, 2013; Accepted March 30, 2013.

Copyright © 2014 Elsevier Inc.

This document may be redistributed and reused, subject to [certain conditions](#).

This document was posted here by permission of the publisher. At the time of the deposit, it included all changes made during peer review, copy editing, and publishing. The U. S. National Library of Medicine is responsible for all links within the document and for incorporating any publisher-supplied amendments or retractions issued subsequently. The published journal article, [guaranteed](#) to be such by Elsevier, is available for free, on ScienceDirect, at: <http://dx.doi.org/10.1016/j.biopsych.2013.03.033>

### Abstract

#### Background

Genome-wide association studies (GWAS) have identified several loci associated with schizophrenia and/or bipolar disorder. We performed a GWAS of psychosis as a broad syndrome rather than within specific diagnostic categories.

#### Methods

1239 cases with schizophrenia, schizoaffective disorder, or psychotic bipolar disorder; 857 of their unaffected relatives, and 2739 healthy controls were genotyped with the Affymetrix 6.0 single nucleotide polymorphism (SNP) array. Analyses of 695,193 SNPs were conducted using UNPHASED, which combines information across families and unrelated individuals. We attempted to replicate signals found in 23 genomic regions using existing data on nonoverlapping samples from the Psychiatric GWAS Consortium and Schizophrenia-GENE-plus cohorts (10,352 schizophrenia patients and 24,474 controls).

#### Results

No individual SNP showed compelling evidence for association with psychosis in our data. However, we observed a trend for association with same risk alleles at loci previously associated with schizophrenia (one-sided  $p = .003$ ). A polygenic score analysis found that the Psychiatric GWAS Consortium's panel of SNPs associated with schizophrenia significantly predicted disease status in our sample ( $p = 5 \times 10^{-14}$ ) and explained approximately 2% of the phenotypic variance.

#### Conclusions

Although narrowly defined phenotypes have their advantages, we believe new loci may also be discovered through meta-analysis across broad phenotypes. The novel statistical methodology we introduced to model effect size heterogeneity between studies should help future GWAS that combine association evidence from related phenotypes. Applying these approaches, we highlight three loci that warrant further investigation. We found that SNPs conveying risk for schizophrenia are also predictive of disease status in our data.

**Key Words:** Bipolar disorder, genome-wide association, meta-analysis, polygenic score analysis,

psychosis, schizophrenia

Psychotic disorders including schizophrenia, bipolar disorders, and schizoaffective disorders affect approximately 3% of the general population (1–6) and constitute the most severe forms of mental diseases. Characteristic symptoms include hallucinations, delusional beliefs, and severe mood variations and cognitive impairments, all of which can lead to major changes in behavior and ability to function. According to the World Health Organization's World Health Report, these psychotic disorders are ranked within the top seven leading causes of disability in young adults (7).

The genetic architecture of schizophrenia and bipolar disorder has been shown to include common alleles of subtle effect and rare mutations of large effect, often involving genome copy number variation (8–11). Recent large-scale meta-analyses of schizophrenia (12), conducted by the Psychiatric GWAS Consortium (PGC), combined data from more than 50,000 individuals from 17 international cohorts (13–25) and identified seven associated loci. Of these loci, five were new, and two had been previously implicated. The strongest new finding in schizophrenia was within an intron of a putative primary transcript for *MIR137* (microRNA 137), a known regulator of neuronal development. Four other schizophrenia loci with strong statistical support contain predicted targets of *MIR137*, suggesting *MIR137*-mediated dysregulation as a previously unknown etiologic mechanism in schizophrenia. The meta-analysis (12) also confirmed the role of the major histocompatibility complex (MHC) region, as suggested in other studies (23,24,26,27), as well as a marker in intron four of transcription factor 4 (*TCF4*) (24).

The PGC conducted a similar meta-analysis for bipolar disorder (28) including more than 11,000 cases and 51,000 controls from previous association studies (15,29–41). The analysis confirmed an association with *CACNA1C* and identified a new intronic variant in *ODZ4*. An overlap in the polygenic component between schizophrenia and bipolar disorder was also found (42,43). In a combined meta-analysis of both schizophrenia and bipolar disorder, three loci reached genome-wide significance: *CACNA1C*, *ANKK3*, and the *ITIH3-ITIH4* region (28).

As data accumulate, there is increasing evidence for overlap in the genetic component to risk between different psychiatric disorders (44–46). When combined with epidemiologic and neuroimaging data (47–50), the shared genetic architecture supports the view of schizophrenia, bipolar disorder, and other psychoses as related rather than etiologically distinct entities (8,12,28,46,51–59). Motivated by these findings, we performed a genome-wide association study (GWAS) of psychotic disorders including patients with schizophrenia, schizoaffective disorder, and bipolar disorder with a history of psychotic symptoms.

## Methods and Materials

### The Cohort

Before any exclusion, the full data set included 6935 participants with 1820 patients, 1224 of their unaffected relatives, and 3891 healthy control subjects. These samples were collected through seven centers across Australia and Europe (Germany, Holland, Spain, and United Kingdom). Participants provided written informed consent, and the study was approved by the respective ethical committees at each of the seven participating centers. After quality control, the full sample included 4835 participants of which 1239, 857, and 2739 were patients, their unaffected relatives, and healthy control subjects, respectively. Additional sample and center details are provided in Table S1 of Supplement 1.

### Inclusion and Exclusion Criteria and Phenotype Definition

To allow for a DSM-IV (60) diagnosis to be ascertained or ruled out, all participants (including controls and unaffected family members) underwent a structured clinical interview with either the Schedule for Affective Disorders and Schizophrenia or the Structured Clinical Interview for DSM Disorders, or the Schedules for Clinical Assessment in Neuropsychiatry (61–63). Of the cases passing quality control, 784 met criteria for schizophrenia, 113 for bipolar disorder with a history of psychotic symptoms, 110 for psychotic disorder not otherwise specified, 97 for schizophreniform disorder, 64 for schizoaffective disorder, 44 for brief psychotic disorder, 20 for delusional disorder, and 7 for substance-induced

psychosis. Participants in all groups were excluded if they had a history of neurologic disease or head injury resulting in loss of consciousness.

### DNA Sample Preparation

Genomic DNA obtained from blood for all participants was sent to the Wellcome Trust Sanger Institute, Cambridge, United Kingdom. Samples were processed in 96-well plate format and each plate carried a positive and a negative control. DNA concentrations were quantified using a PicoGreen assay (Invitrogen, Life Technologies, Grand Island, New York) and an aliquot assayed by agarose gel electrophoresis. A sample passed quality control if the original DNA concentration was at least 50 ng/ $\mu$ L and the DNA was not degraded.

### Genotyping Methodology and Quality Control

To track sample identity, 30 single nucleotide polymorphisms (SNPs) including sex chromosome markers were typed on the Sequenom platform before entry to the whole genome genotyping pipeline. Of the initial 6935 samples, 347 failed quality control due to degraded or insufficient DNA or incorrect sex classification. The remaining samples were sent for genotyping with the Genome-wide Human SNP Array 6.0 at Affymetrix Services Lab (<http://www.affymetrix.com>).

### Data Quality Control

See [Tables S2](#) and [S3](#) in [Supplement 1](#). Genotype calling was conducted using the CHIAMO algorithm ([64,65](#)) modified for use with the Affymetrix 6.0 genotyping array. We excluded 11,610 SNPs with a study-wide missing data rate over 5%. We removed 26,858 SNPs with four or more Mendelian inheritance errors identified with Pedstats ([66](#)). Additional exclusion criteria were departure from Hardy-Weinberg equilibrium ( $p < 10^{-6}$ ) or minor allele frequency (MAF)  $< .02$  with 2404 and 145,097 SNPs removed, respectively. A total of 38,895 SNPs from the X or Y chromosomes or mitochondrial DNA were also excluded from the analysis. Finally, 9499 poorly genotyped SNPs were removed following visual inspection of the genotyping intensity plots in the program Evoker ([67](#)).

We excluded 214 samples with more than 2% missing data across all SNPs. Another 70 samples were excluded due to divergent genome-wide heterozygosity (inbreeding coefficients were  $F > .076$  or  $F < -.076$  as estimated with PLINK ([42](#)). Chromosomal sharing was inferred from a genome-wide subset of 71,677 SNPs and from each duplicate pair the sample with the most complete genotype data was kept. We removed 70 duplicates and monozygotic twins by excluding one of each pair of individuals showing identity by descent greater than 95%.

After quality control, 4835 individuals remained. Initial analysis of the genotype data identified a high fraction of samples (approximately 30%), which showed poor signal-to-noise ratio in the genotyping assay. Because the experimental source of the problem was unclear and to ensure a robust set of genotype calls, these samples were removed from further analysis. We note that the sample loss was randomly distributed across the three clinical groups (32% of patients, 30% of relatives, and 30% of controls;  $\chi^2$  (2 df) = 3.2;  $p = .20$ ). Full details on the sample quality control are provided in [Table S2](#) of [Supplement 1](#).

In addition to 3490 unrelated individuals, there were 1345 related individuals clustered in 462 families. The family size ranged between 2 and 5 with an average of 2.9 members. Of the families, 196 were control families, 243 had one affected case only, 21 families included two cases, and another 2 families had three cases. Data from these individuals were analyzed at 695,193 autosomal SNPs.

### Population Structure Analysis

To investigate the genetic structure in the data, we performed principal component (PC) analysis (PCA) of unrelated individuals using EIGENSOFT version 3.0 ([68](#)) on a thinned set of SNPs (see Methods and Materials in [Supplement 1](#)). Owing to the multicenter nature of our study, we assessed the need to include PCs as covariates in statistical tests of association to control for population stratification ([69](#)). This was done by using PLINK ([42](#)) to calculate the genome-wide distribution of the association test statistic in the unrelated individuals using different numbers of PCs as covariates. Possible inflation in the test statistic was measured by the genomic control parameter  $\lambda$ , which is the ratio of the median of

the observed test statistic distribution to that of its expectation under the null hypothesis (70).

### Association Analysis in Our Discovery Sample

A genome-wide association analysis was conducted with UNPHASED v3.1.4. (71), which allows a combined analysis of both families and unrelated individuals, thus increasing statistical power. UNPHASED calculates separately the transmitted and untransmitted alleles in families as well as the allelic frequencies in unrelated patients and controls, giving a combined odds ratio, 95% confidence interval, and  $p$  value. The analysis included three PCs as covariates.

For SNPs showing association with psychosis in our data with  $p < 1 \times 10^{-4}$ , proxy SNPs were identified using the proxy report routine in PLINK (42). Only those SNPs that were in linkage disequilibrium ( $r^2 \geq .5$ ) with and within 100-kb distance from at least one such proxy SNP that showed association with psychosis with  $p < 1 \times 10^{-2}$  were selected for the replication phase. These criteria reduced the possibility that the association signal was driven by an artifact at the most associated SNP.

We attempted to replicate 44 SNPs included in the catalogue of published GWAS (72), accessed in January 2012, for schizophrenia or bipolar disorder with  $p$  values less than  $1 \times 10^{-7}$ . These SNPs and the studies that identified them are listed in Table 1. If a reported SNP was not genotyped directly in our data, we used the 1000 Genomes Project (73) data to identify the best tag (highest  $r^2$ ) and orientated the haplotype to the risk allele so that the directions of the odds ratios were matched between our analysis and the previous studies. Where relevant information was not available, the SNP was excluded from analysis.

### Related Effects Meta-Analysis

We attempted in silico replication with independent samples from the PGC and the SGENE+ consortia using both a fixed effects meta-analysis (42) and a Bayesian related effects approach (27). The Bayesian related effects model (see Methods and Materials in Supplement 1) was considered appropriate in this context because there are obvious sources of heterogeneity between the discovery and the replication datasets due to differing phenotype definitions and sample origins. To ensure that no samples were overlapping between the discovery and replication data sets, we used a subset of at least 13,000 SNPs to exclude 717 participants (51 cases, 642 controls, 24 relatives) from our primary study. Similarly, another 74 control subjects were excluded from the SGENE+ replication cohort. Only SGENE+ samples that were not part of the PGC study were included in our meta-analyses.

### Polygenic Score Analysis

A polygenic score analysis was conducted using a thinned (on the basis of linkage disequilibrium) panel of 113,774 SNPs from the PGC-schizophrenia meta-analysis. Of these SNPs, 72,635 were available in our data, and we further added suitable proxies for another 19,434 SNPs, which were identified through the program SNAP (75) to have  $r^2 > .9$  with the index SNP and to be within 500-kb distance. Of 4835 samples passing the above described quality control filters we excluded 717 individuals that overlapped with the PGC sample and another 833 related participants. Therefore, our polygenic analysis was based on a subset of 3285 unrelated subjects comprising of 1188 patients and 2097 controls. The polygenic score for each individual was calculated from the number of risk alleles they carried for each SNP, weighted by the log(OR) provided by the PGC, and summed across all the SNPs. Calculations were performed using the PLINK SNP scoring routine (42).

We used logistic regression, with the three population structure PCs and the center of origin of the samples as covariates, to test whether the polygenic scores were predictive of case-control status in our study. Following Purcell *et al.* (26), we reported the proportion of the variance of disease risk in our sample (as measured by Nagelkerke's pseudo- $R^2$ ) that could be explained by the panel of SNPs. We examined the proportion of the variance explained by our data at various  $p$  value thresholds for the PGC-schizophrenia SNP list ( $p < .01, .05, .1, .2, .3, .4$ , and  $.5$ ) (26,43).

## Results

The primary goal of our study was to use published GWAS data in schizophrenia and bipolar disorder to facilitate discovery in a new sample with a broad psychosis phenotype. First, we examine previously



published loci in our data set. We then present our discovery association data followed by replication in independent samples, using a novel meta-analysis approach. Finally, we report a polygenic score analysis testing whether a panel of SNPs associated specifically with schizophrenia are predictive of the broader phenotype in our sample.

### Population Structure

Analysis of our data via PCA showed that the major components of the genetic variation capture the geographic origin of the samples ([Figure S1](#) in [Supplement 1](#)). The first PC separates samples from northern Spain (Santander) from the rest of the sample. We found it necessary to include the first three PCs as covariates to control for the confounding effects of population structure even though this may reduce power to detect associations. The genomic control parameter,  $\lambda$ , was equal to 1.030 and 1.374 with and without including the three PCs respectively.

### Replication of Previously Published Loci in Our Data Set

To ascertain the strength of association in our data at loci previously associated with schizophrenia and/or bipolar disorder, we systematically assessed the evidence at the SNPs most strongly correlated with the previously reported SNPs and their risk alleles ([Table 1](#); [Figure 1](#)) ([12,15,21,23,24,26,28,30,33,34,37,38,66,76–96](#)). We observed evidence for association at several loci. These include multiple SNPs in the MHC region, most strongly at rs9272219 near *HLA-DQA1*. We also replicated association at *MIR137*, *MCTP1*, *NRGN*, and *CDC68*. More generally, we observed a positive correlation between previously reported effect sizes and those estimated here, most notably with schizophrenia associated loci where 19 of the 24 loci had the same risk allele (one-sided  $p = .003$ ). We note that this sign test remained significant ( $p < .05$ ) after thinning to reduce correlation between SNPs and removing the SNPs within the MHC ([Table S4](#) in [Supplement 1](#)).

### Association Analysis in Our Discovery Sample

We did not identify any novel SNPs associated with psychosis as a broad phenotype with  $p < 1 \times 10^{-7}$  from our data set. [Figure 2](#) and [Figure S2](#) in [Supplement 1](#) show respectively the quantile-quantile plots of the distribution of observed and expected  $p$  values and the Manhattan plot of all autosomal SNPs that passed quality control.

### Independent Replication of Our Discovery Data

We selected 63 SNPs in 23 genomic regions ([Supplementary Data](#) in [Supplement 1](#)) showing the strongest evidence for association in our discovery dataset ( $p < 1 \times 10^{-4}$ ) and having at least one suitable proxy SNP also associated (see Methods and Materials). We then attempted in silico replication using the large independent (see Methods and Materials) data sets publicly available from the PGC ([51,57,97,98](#)) as well as with data provided by the SGENE+ Consortium ([24,25](#)).

All of the selected 63 SNPs were present in at least one of the replication data sets, and 53 SNPs were available in both collections giving a two-study independent replication data set with 10,352 patients and 24,474 control subjects. By combining estimates of the logarithmic odds ratios weighted by their precision (i.e., inverse variance of the estimator) across discovery and replication cohorts, we conducted a three-study meta-analysis with 11,540 patients, 833 of their unaffected relatives, and 26,571 control subjects. [Table 2](#) contains the SNPs that in the replication data showed an effect in the same direction (with one-sided  $p < .05$ ) as in the discovery data set and had either combined (discovery plus replication) fixed effects  $p$  value  $< 10^{-3}$  or combined fixed effects or related effects Bayes factor  $> 100$ . Regional association plots of these SNPs are included in [Figure 3](#) and forest plots are given in [Figure S3](#) in [Supplement 1](#). Full details of the replication meta-analyses are presented in [Supplement 1](#).

Two of five top regions in our discovery analysis, rs743393 (on chromosome 3) and rs968794 (on chromosome 7), showed effects in the same direction in the replication data with one-sided replication  $p$  values of .020 and .007, respectively. However, the combined fixed effects evidence of association at these loci is weak by the standards of GWAS ( $p$  values  $7 \times 10^{-4}$  and  $8 \times 10^{-5}$ , and Bayes factors 25 and 200, respectively). Bayes factors from the related effects model (440 and 7400, respectively), which allow heterogeneity in effect size across studies, are larger than those from the fixed effect model but

still fail to provide convincing levels of evidence for these associations. See Wellcome Trust Case Control Consortium ([66](#)) for interpreting Bayes factors in the GWAS context.

### Polygenic Score Analysis

We performed a polygenic score analysis using the SNPs associated with schizophrenia in the PGC study. Logistic regression analyses showed significantly higher scores for patients than control subjects ( $p = 5 \times 10^{-14}$ ), and explained approximately 1.7% of Nagelkerke's pseudo-variance in our sample. For additional results on the polygenic component, see Methods and Materials and [Figure S4](#) in [Supplement 1](#).

### Discussion

Although extensive twin research has shown high heritability estimates ranging between 73% and 85% for schizophrenia, schizoaffective disorder, and manic syndromes ([99–103](#)), compared with other neuropsychiatric diseases, the search for psychosis susceptibility genes has not been straightforward ([104–110](#)). Several loci have now been convincingly identified, and more are expected to be found through the PGC, a large international collaboration undertaking mega-analyses of the available data ([8,12,28](#)). In this context, our sample with 4835 participants has modest power to detect new variants unequivocally ([111](#)), and thus it is not that surprising that we did not identify convincing novel associations with a broad psychosis phenotype. For example, our data provide less than 10% power at type I error of  $1 \times 10^{-7}$  for an allelic relative risk of 1.2 at a common SNP, and in practice power may be further reduced because of our broader disease phenotype and our use of PC covariates to control for population structure. It has been estimated that for allelic odds ratios of 1.1 and 1.2 as many as 10,000 to more than 20,000 subjects would be required to achieve 80% power at genome-wide significance levels ([112](#)). We hope that our data set will be valuable in the future meta-analyses that are likely to identify further common variants for psychotic illnesses with convincing statistical evidence.

The loci for which we found the best evidence of association, given our data and the independent replication samples from the PGC and SGENE+ consortium, are reported in [Table 2](#) and their regional association plots are in [Figure 3](#). These include rs743393 located in chromosome 3, which is near the Integrin alpha 9 gene (ITGA9), thought to codify for a membrane glycoprotein that mediates cell-cell and cell-matrix adhesion ([22](#)). SNP rs968794 in chromosome 7 is close to the NDUFA4 gene coding for a protein that is involved in the mitochondrial membrane respiratory chain ([23](#)). Finally, rs4761708 is located in a gene desert in chromosome 12 flanked by recombination hotspots. Limited information is available on the role of these genes and whether they are causal in increasing the risk of developing psychosis. Additional data are required to confirm the associations and to inform on underlying disease mechanisms.

Although our sample was too small to identify new loci unequivocally, it had good power to replicate previously published signals with modest effects assuming they are consistent across psychosis phenotypes. The analysis of established schizophrenia loci shows that there is a trend for consistent association in our data ([Figure 1](#)). Also the polygenic score analysis showed that, compared with control subjects, our cases are enriched for alleles thought to convey risk for schizophrenia; the PGC's panel of SNPs significantly predicted case–control status in our study and explained around 2% of the phenotypic variance in our sample. This is comparable to previously reported figures in other European case–control and family-based populations ([26,43](#)) and is consistent with a highly polygenic model of disease risk. Larger data sets are required to determine whether these signals are specific to only some of the subphenotypes of psychosis or are present across them.

Schizophrenia and bipolar disorder are thought to be highly complex and polygenic, with potentially thousands of susceptibility genes of small effect and, particularly for schizophrenia, also with rare variants of larger effects ([10,14,26,52,58,107,110–113](#)). Previous GWAS have generally targeted a specific diagnostic category, and the merit of analyzing a psychosis spectrum is debatable. One of the challenges facing psychiatric genetics has been the heterogeneity of the diseases in question and studying a broader phenotype could exacerbate this problem. However, as GWAS data accumulate, there is growing evidence of genetic overlap between schizophrenia and bipolar disorder ([8,12,28,44,46,51,55,56,58–60](#)). Therefore, in addition to examining narrowly defined diseases it is of

use to undertake cross-disorder analyses. Our replication of established schizophrenia hits in a broadly defined psychosis sample and polygenic analyses support this view.

Current benchmarks for establishing new associations usually insist on a combined discovery and replication  $p$  value of less than  $5 \times 10^{-8}$  using a fixed effects meta-analysis, which is the correct summary of the overall statistical evidence when the true effect size is exactly the same across the studies. However, when the effect size may genuinely differ between the studies due to, for example, differences in phenotype definition, patterns of linkage disequilibrium or environmental contributions, it is appropriate to assess association evidence with a model that allows for heterogeneity. Here we have used a Bayesian related effects model that assumes that the effects between the studies are similar (e.g., highly correlated on the log odds scale) but not necessarily the same. A computational advantage of this model is that it only requires the studywide summary statistics (effect size and its standard error), and therefore we believe it will be useful in future meta-analyses of existing GWAS data sets, for example, across neuropsychiatric diseases.

In our study, we applied the related effects model to combine the discovery and replication results at 63 SNPs that had the most evidence in the discovery analysis. As expected, the related effects model assigns a higher probability than the fixed effects model to the event that these SNPs represent real associations ([Supplement 1](#)). For any one SNP, this can be either because the association is a true positive but heterogeneous and/or subject to a “winner’s curse” effect ([114,115](#)) or because the association is a false positive and appears heterogeneous because of a relatively strong discovery signal created by chance effects. Future studies are needed to determine conclusively whether the SNPs reported in [Table 2](#) and [Figure 3](#) represent real associations.

The scenario described here in which a discovery analysis of a particular phenotype in a small sample attempts to replicate putative associations in silico, in much larger meta-analyses of similar phenotypes, is likely to become more common in human genetics as GWAS data accumulate. We note that stringent  $p$  value thresholds for combined discovery and replication meta-analysis make it challenging to identify new loci because loci with consistent effects across studies will have already been identified. Therefore, an ongoing genome-wide meta-analysis across related phenotypes modeling effect heterogeneity where appropriate is arguably the most powerful approach to identify additional susceptibility loci.

In conclusion, although no individual SNPs reached convincing evidence of association with psychosis, our data contribute to the international effort to produce a cohort large enough to investigate both the shared and separate genetic bases of multiple psychotic disorders. Our polygenic score analysis indicates that the PGC-derived panel of SNPs conveying risk specifically for schizophrenia is also predictive of case–control status in our data, explaining around 2% of the phenotypic variance. We advocate the use of related effects models in large meta-analyses for summarizing the evidence for association across independent studies of related phenotypes.

### The authors of this paper are as follows:

Elvira Bramon<sup>1,2\*</sup>, Matti Pirinen<sup>3\*</sup>, Amy Strange<sup>3</sup>, Kuang Lin<sup>2</sup>, Colin Freeman<sup>3</sup>, Céline Bellenguez<sup>3</sup>, Zhan Su<sup>3</sup>, Gavin Band<sup>3</sup>, Richard Pearson<sup>3</sup>, Damjan Vukcevic<sup>3</sup>, Cordelia Langford<sup>4</sup>, Panos Deloukas<sup>4</sup>, Sarah Hunt<sup>4</sup>, Emma Gray<sup>4</sup>, Serge Dronov<sup>4</sup>, Simon C. Potter<sup>4</sup>, Avazeh Tashakkori-Ghanbaria<sup>4</sup>, Sarah Edkins<sup>4</sup>, Suzannah J. Bumpstead<sup>4</sup>, Maria J. Arranz<sup>2,5</sup>, Steven Bakker<sup>6</sup>, Stephan Bender<sup>7,8</sup>, Richard Bruggeman<sup>9</sup>, Wiepke Cahn<sup>6</sup>, David Chandler<sup>10</sup>, David A Collier<sup>2,11</sup>, Benedicto Crespo-Facorro<sup>12,13</sup>, Paola Dazzan<sup>2</sup>, Lieuwe de Haan<sup>14</sup>, Marta Di Forti<sup>2</sup>, Milan Dragović<sup>15</sup>, Ina Giegling<sup>16</sup>, Jeremy Hall<sup>17</sup>, Conrad Iyegbe<sup>2</sup>, Assen Jablensky<sup>15</sup>, René S. Kahn<sup>6</sup>, Luba Kalaydjieva<sup>10</sup>, Eugenia Kravariti<sup>2</sup>, Stephen Lawrie<sup>17</sup>, Don H. Linszen<sup>14,19</sup>, Ignacio Mata<sup>12,13</sup>, Colm McDonald<sup>18</sup>, Andrew McIntosh<sup>17</sup>, Inez Myin-Germeys<sup>19</sup>, Roel A. Ophoff<sup>20,6</sup>, Carmine M. Pariante<sup>2</sup>, Tiina Paunio<sup>21,22</sup>, Marco Picchioni<sup>2,23</sup>, Psychiatric Genomics Consortium<sup>24</sup>, Stephan Ripke<sup>25,26</sup>, Dan Rujescu<sup>16,27</sup>, Heinrich Sauer<sup>28</sup>, Madiha Shaikh<sup>2,29</sup>, Jessika Sussmann<sup>17</sup>, Jaana Suvisaari<sup>30</sup>, Sarah Tosato<sup>31</sup>, Timothea Touloupoulou<sup>2,32,33</sup>, Jim Van Os<sup>2,19</sup>, Muriel Walshe<sup>2</sup>, Matthias Weisbrod<sup>8</sup>, Heather Whalley<sup>17</sup>, Durk Wiersma<sup>9</sup>, Jenefer M. Blackwell<sup>34,35</sup>, Matthew A. Brown<sup>36</sup>, Juan P. Casas<sup>37,38</sup>, Aiden Corvin<sup>39</sup>, Audrey Duncanson<sup>40</sup>, Janusz A.Z. Jankowski<sup>41,42,43</sup>, Hugh S. Markus<sup>44</sup>, Christopher G. Mathew<sup>45</sup>, Colin N.A. Palmer<sup>46</sup>, Robert Plomin<sup>47</sup>, Anna Rautanen<sup>3</sup>, Stephen J. Sawcer<sup>48</sup>, Richard C. Trembath<sup>44</sup>, Nicholas W. Wood<sup>49</sup>, Ines Barroso<sup>4</sup>, Leena Peltonen<sup>4,‡</sup>, Cathryn M. Lewis<sup>2,46</sup>, Robin M. Murray<sup>2</sup>, Peter Donnelly<sup>3,50</sup>, John Powell<sup>2</sup>, ^ and Chris C.A. Spencer<sup>3</sup>, ^\*These authors contributed equally to this work.^These authors jointly supervised this work.‡Deceased.Author affiliations:



<sup>1</sup>Mental Health Sciences Unit & Institute of Cognitive Neuroscience, University College London, UK. <sup>2</sup>Institute of Psychiatry – King’s College London, De Crespigny Park, London, UK. <sup>3</sup>Wellcome Trust Centre for Human Genetics, University of Oxford, UK. <sup>4</sup>Wellcome Trust Sanger Institute, Cambridge, UK. <sup>5</sup>Fundació de Docència i Recerca Mútua de Terrassa, Universitat de Barcelona, Catalonia, Spain. <sup>6</sup>University Medical Center Utrecht, Department of Psychiatry, Rudolf Magnus Institute of Neuroscience, The Netherlands. <sup>7</sup>Child and Adolescent Psychiatry, University of Technology Dresden, Dresden, Germany. <sup>8</sup>Section for Experimental Psychopathology and Neurophysiology, General Psychiatry, Heidelberg, Germany. <sup>9</sup>University Medical Center Groningen, Department of Psychiatry, University of Groningen, The Netherlands. <sup>10</sup>Western Australian Institute for Medical Research and Centre for Medical Research, The University of Western Australia, Perth, Australia. <sup>11</sup>Discovery Neuroscience Research, Eli Lilly and Company Ltd, UK. <sup>12</sup>University Hospital Marqués de Valdecilla, IFIMAV, Department of Psychiatry, School of Medicine, University of Cantabria, Santander, Spain. <sup>13</sup>CIBERSAM, Centro Investigación Biomédica en Red Salud Mental, Madrid, Spain. <sup>14</sup>Academic Medical Centre University of Amsterdam, Department of Psychiatry, Amsterdam, The Netherlands. <sup>15</sup>Centre for Clinical Research in Neuropsychiatry, The University of Western Australia, Perth, Australia. <sup>16</sup>University of Munich, Department of Psychiatry, Munich, Germany. <sup>17</sup>Division of Psychiatry, Kennedy Tower, University of Edinburgh, Edinburgh, Scotland, UK. <sup>18</sup>Department of Psychiatry, Clinical Science Institute, National University of Ireland, Galway, Ireland. <sup>19</sup>Maastricht University Medical Centre, South Limburg Mental Health Research and Teaching Network, EURON, Maastricht, The Netherlands. <sup>20</sup>University of California Los Angeles Center for Neurobehavioral Genetics, Los Angeles, California, USA. <sup>21</sup>National Institute for Health and Welfare, Public Health Genomics Unit, Helsinki, Finland. <sup>22</sup>Department of Psychiatry, University of Helsinki, Helsinki, Finland. <sup>23</sup>St Andrews Academic Centre, Institute of Psychiatry, Northampton, UK. <sup>24</sup>Psychiatric Genomics Consortium. <https://pgc.unc.edu/>. A list of members is provided in the Supplement. <sup>25</sup>Analytic and Translational Genetics Unit, Massachusetts General Hospital, Harvard Medical School, Boston, Massachusetts, USA. <sup>26</sup>Broad Institute of MIT and Harvard University, Cambridge, Massachusetts, USA. <sup>27</sup>University of Halle, Department of Psychiatry, Halle, Germany. <sup>28</sup>Department of Psychiatry and Psychotherapy, University Hospital of Jena, Jena, Germany. <sup>29</sup>Department of Psychology, Royal Holloway, University of London, Egham, Surrey, UK. <sup>30</sup>National Institute for Health and Welfare, Department of Mental Health and Substance Abuse Services, Helsinki, Finland. <sup>31</sup>Section of Psychiatry, University of Verona, Verona, Italy. <sup>32</sup>Department of Psychology, The University of Hong Kong, Hong Kong, SAR of the People’s Republic of China. <sup>33</sup>State Key Laboratory of Brain and Cognitive Sciences, The University of Hong Kong, SAR of the People’s Republic of China. <sup>34</sup>Telethon Institute for Child Health Research, Centre for Child Health Research, University of Western Australia, Subiaco, Western Australia. <sup>35</sup>Cambridge Institute for Medical Research, University of Cambridge School of Clinical Medicine, Cambridge, UK. <sup>36</sup>University of Queensland, Diamantina Institute, Princess Alexandra Hospital, University of Queensland, Brisbane, Queensland, Australia. <sup>37</sup>Department of Epidemiology and Population, Health London School of Hygiene and Tropical Medicine, London, UK. <sup>38</sup>Department of Epidemiology and Public Health, University College London, UK. <sup>39</sup>Neuropsychiatric Genetics Research Group, Institute of Molecular Medicine, Trinity College Dublin, Ireland. <sup>40</sup>Molecular and Physiological Sciences, The Wellcome Trust, UK. <sup>41</sup>Centre for Digestive Diseases, Blizard Institute Queen Mary University of London, UK. <sup>42</sup>Department of Medical Oncology, Churchill Hospital University of Oxford, UK. <sup>43</sup>Digestive Disease Academic Centre, Leicester Royal Infirmary, Leicester, UK. <sup>44</sup>Stroke and Dementia Research Centre, St George’s University of London, UK. <sup>45</sup>King’s College London, Department Medical and Molecular Genetics King’s Health Partners Guy’s Hospital, London, UK. <sup>46</sup>Biomedical Research Centre, Ninewells Hospital and Medical School, Dundee, UK. <sup>47</sup>King’s College London, Social Genetic and Developmental Psychiatry Centre Institute of Psychiatry, London, UK. <sup>48</sup>University of Cambridge, Department Clinical Neurosciences, Addenbrooke’s Hospital Cambridge, UK. <sup>49</sup>Department Molecular Neuroscience, Institute of Neurology, Queen Square London, UK. <sup>50</sup>Department of Statistics, University of Oxford, Oxford, UK.

## Acknowledgments

The authors thank all the patients, relatives and controls who took part in this research as well as the clinical staff who facilitated their involvement.

*The principal funding for this study was provided by the Wellcome Trust, as part of the Wellcome Trust Case Control Consortium 2 project (Grant Nos. 085475/B/08/Z and 085475/Z/08/Z).*

*E. Bramon currently holds a MRC New Investigator Award and a MRC Centenary Award. E. Bramon was further supported by the National Institute of Health Research UK (post-doctoral fellowship), the Psychiatry Research Trust, the Schizophrenia Research Fund, the Brain and Behavior Research foundation's (NARSAD's) Young Investigator Award, a Wellcome Trust Research Training Fellowship and the NIHR Biomedical Research Centre for Mental Health at the South London and Maudsley NHS Foundation Trust and Institute of Psychiatry Kings College London.*

*P. Donnelly was supported in part by a Royal Society Wolfson Merit Award. C. Spencer was supported by a Wellcome Trust Fellowship (097364/Z/11/Z). M. Pirinen is supported by the Academy of Finland (257654). We acknowledge use of the British 1958 Birth Cohort DNA collection funded by the Medical Research Council (grant G0000934) and the Wellcome Trust (grant 068545/Z/02), the UK National Blood Service controls funded by the Wellcome Trust and the People of the British Isles collection, funded by the Wellcome Trust. Work was supported in part by Wellcome Trust Centre for Human Genetics core grants 072894/Z/03/Z, 090532/Z/09/Z and 075491/Z/04/B.*

*M. Picchioni acknowledges a Wellcome Trust Research Training Fellowship (064971). T. Touloupoulou acknowledges the European Community's Sixth Framework Programme through a Marie Curie Training Network (MRTN-CT-2006-035987; Local Co-PIs Murray and Touloupoulou) called the European Twin Study Network on Schizophrenia (EUTwinsS), NARSAD (through a Young Investigator Award), the Economic and Social Research Council/Medical Research Council, the Psychiatry Research Trust (PTA-037-27-0002) and the NIHR Biomedical Research Centre for Mental Health at the South London and Maudsley NHS Foundation Trust and Institute of Psychiatry Kings College London. All co-authors based at the Institute of Psychiatry—King's College London acknowledge the NIHR Biomedical Research Centre for Mental Health at the South London and Maudsley NHS Foundation Trust and Institute of Psychiatry Kings College London.*

*The Edinburgh cohort was supported by: a NARSAD independent investigator award and a Health Foundation Clinician Scientist Award (to A. McIntosh), a Wellcome Trust Clinical Research Training Fellowship (to Jessika Sussman) and a Royal Society Dorothy Hodgkin Fellowship (to Heather Whalley).*

*The GROUP study was supported by: (i) The Geestkracht programme of the Dutch Health Research Council (ZON-MW, grant number 10-000-1002). (ii) The EU Seventh Framework Programme (consortium name: EU-GEI; No. HEALTH-F2-2009-241909) and matching funds from participating universities and mental health care organizations (Site Amsterdam: Academic Psychiatric Centre AMC, Ingeest, Arkin, Dijk en Duin, Rivierduinen, Erasmus MC, GGZ Noord Holland Noord; Site Utrecht: University Medical Centre Utrecht, Altrecht, Symfora, Meerkanten, Riagg Amersfoort, Delta; Site Groningen: University Medical Center Groningen, Lentis, GGZ Friesland, GGZ Drenthe, Adhesie, Mediant, GGZ De Grote Rivieren and Parnassia psycho-medical centre; Site Maastricht: Maastricht University Medical Center, GGZ Eindhoven, GGZ Midden-Brabant, GGZ Oost-Brabant, GGZ Noord-Midden Limburg, Mondriaan Zorggroep, Prins Clauscentrum Sittard, RIAGG Roermond, Universitair Centrum Sint-Jozef Kortenberg, CAPRI University of Antwerp, PC Ziekeren Sint-Truiden, PZ Sancta Maria Sint-Truiden, GGZ Overpelt, OPZ Rekem). (iii) NIMH grant to R.A. Ophoff (1R01MH078075).*

*The Heidelberg cohort work was supported by a grant (WE-1996/1-3) from the German Research Community (Deutsche Forschungsgemeinschaft, DFG) to M. Weisbrod. We would like to thank Dr. Anuradha Sharma for her help with the management of the Heidelberg dataset.*

*The Perth cohort was supported by grant 513874 of the National Health and Medical Research Council, Australia and by the North Metropolitan Health Service, Perth, Australia.*

*The Santander cohort was supported by Instituto de Salud Carlos III (PI020499, PI050427, PI060507), SENY Fundació (CI 2005-0308007), Fundacion Ramón Areces and Fundacion Marqués de Valdecilla (API07/011, API10/13).*

*We thank S. Bertrand, J. Bryant, S.L. Clark, J.S. Conquer, T. Dibling, J.C. Eldred, S. Gamble, C. Hind, M.L. Perez, C.R. Stribling, S. Taylor and A. Wilk of the Wellcome Trust Sanger Institute's Sample and*

*Genotyping Facilities for technical assistance. We thank Engilbert Sigurdsson, Kari Stefansson, Hreinn Stefansson and Stacy Steinberg from the SGENE plus consortium for contributing to the replication analyses. We thank the Psychiatric Genomics Consortium (<http://pgc.unc.edu>) for assistance with replication and polygenic score analyses.*

#### Disclosures

*B. Crespo-Facorro has received honoraria for his participation as a speaker at educational events from Bristol-Myers Squibb and Johnson & Johnson. J. Jankowski is a consultant to Astra Zeneca. S. Lawrie, A. McIntosh, J. Hall and H. Whalley have all received support from Pfizer (formerly Wyeth) for imaging studies of schizophrenia and bipolar disorder. M. Picchioni receives financial support from Pfizer and Janssen Cilag. J. Van Os has received unrestricted investigator-led research grants or recompense for presenting his research from Eli Lilly, Bristol-Myers Squibb, Lundbeck, Organon, Janssen-Cilag, GlaxoSmithKline, AstraZeneca, Pfizer and Servier, companies that have an interest in the treatment of psychosis. All other co-authors declare no biomedical financial interests or potential conflict of interests.*

#### Membership of the Psychosis Endophenotypes International Consortium (PEIC):

Maria J. Arranz<sup>1,2</sup>, Steven Bakker<sup>3</sup>, Stephan Bender<sup>4,5</sup>, Elvira Bramon<sup>6,2</sup>, David Collier<sup>7,2</sup>, Benedicto Crespo-Facorro<sup>8,9</sup>, Jeremy Hall<sup>10</sup>, Conrad Iyegbe<sup>2</sup>, Assen Jablensky<sup>11</sup>, René S. Kahn<sup>3</sup>, Luba Kalaydjieva<sup>12</sup>, Stephen Lawrie<sup>10</sup>, Cathryn M. Lewis<sup>2</sup>, Kuang Lin<sup>2</sup>, Don H. Linszen<sup>13</sup>, Ignacio Mata<sup>8,9</sup>, Andrew McIntosh<sup>10</sup>, Robin M. Murray<sup>2</sup>, Roel A. Ophoff<sup>14</sup>, John Powell<sup>2</sup>, Dan Rujescu<sup>15,16</sup>, Jim Van Os<sup>17,2</sup>, Muriel Walshe<sup>2</sup>, Matthias Weisbrod<sup>5</sup> and Durk Wiersma<sup>18</sup>

#### PEIC Affiliations:

<sup>1</sup>Fundació de Docència i Recerca Mútua de Terrassa, Universitat de Barcelona, Catalonia, Spain.

<sup>2</sup>Institute of Psychiatry at King's College London, De Crespigny Park, London, UK. <sup>3</sup>University Medical Center Utrecht, Department of Psychiatry, Rudolf Magnus Institute of Neuroscience, The Netherlands.

<sup>4</sup>Child and Adolescent Psychiatry, University of Technology Dresden, Dresden, Germany. <sup>5</sup>Section for Experimental Psychopathology, General Psychiatry, Heidelberg, Germany. <sup>6</sup>Mental Health Sciences Unit & Institute of Cognitive Neuroscience, University College London, UK. <sup>7</sup>Discovery Neuroscience Research, Lilly, UK. <sup>8</sup>University Hospital Marqués de Valdecilla, IFI MAV, Department of Psychiatry, School of Medicine, University of Cantabria, Santander, Spain. <sup>9</sup>CIBERSAM, Centro Investigación Biomédica en Red Salud Mental, Madrid, Spain. <sup>10</sup>Division of Psychiatry, Kennedy Tower, University of Edinburgh, Edinburgh, Scotland, UK. <sup>11</sup>Centre for Clinical Research in Neuropsychiatry, The University of Western Australia, Perth, Australia. <sup>12</sup>Western Australian Institute for Medical Research and Centre for Medical Research, The University of Western Australia, Perth, Australia. <sup>13</sup>Academic Medical Centre University of Amsterdam, Department of Psychiatry, Amsterdam, The Netherlands.

<sup>14</sup>UCLA Center for Neurobehavioral Genetics, Los Angeles, CA, USA. <sup>15</sup>University of Munich, Dept. of Psychiatry, Munich, Germany. <sup>16</sup>University of Halle, Dept. of Psychiatry, Halle, Germany.

<sup>17</sup>Maastricht University Medical Centre, South Limburg Mental Health Research and Teaching Network, EURON, Maastricht, The Netherlands. <sup>18</sup>University Medical Center Groningen, Department of Psychiatry, University of Groningen, The Netherlands.

<sup>14</sup>UCLA Center for Neurobehavioral Genetics, Los Angeles, CA, USA. <sup>15</sup>University of Munich, Dept. of Psychiatry, Munich, Germany. <sup>16</sup>University of Halle, Dept. of Psychiatry, Halle, Germany.

<sup>17</sup>Maastricht University Medical Centre, South Limburg Mental Health Research and Teaching Network, EURON, Maastricht, The Netherlands. <sup>18</sup>University Medical Center Groningen, Department of Psychiatry, University of Groningen, The Netherlands.

<sup>14</sup>UCLA Center for Neurobehavioral Genetics, Los Angeles, CA, USA. <sup>15</sup>University of Munich, Dept. of Psychiatry, Munich, Germany. <sup>16</sup>University of Halle, Dept. of Psychiatry, Halle, Germany.

#### Membership of Wellcome Trust Case Control Consortium 2 (WTCCC2):

Management Committee: Peter Donnelly (Chair)<sup>1,2</sup>, Ines Barroso (Deputy Chair)<sup>3</sup>, Jenefer M. Blackwell<sup>4,5</sup>, Elvira Bramon<sup>6</sup>, Matthew A. Brown<sup>7</sup>, Juan P. Casas<sup>8</sup>, Aiden Corvin<sup>9</sup>, Panos Deloukas<sup>3</sup>, Audrey Duncanson<sup>10</sup>, Janusz Jankowski<sup>11</sup>, Hugh S. Markus<sup>12</sup>, Christopher G. Mathew<sup>13</sup>, Colin N.A. Palmer<sup>14</sup>, Robert Plomin<sup>15</sup>, Anna Rautanen<sup>1</sup>, Stephen J. Sawcer<sup>16</sup>, Richard C. Trembath<sup>13</sup>, Ananth C. Viswanathan<sup>17</sup>, Nicholas W. Wood<sup>18</sup>

Data and Analysis Group: Chris C.A. Spencer<sup>1</sup>, Gavin Band<sup>1</sup>, Céline Bellenguez<sup>1</sup>, Colin Freeman<sup>1</sup>, Garrett Hellenthal<sup>1</sup>, Eleni Giannoulatou<sup>1</sup>, Matti Pirinen<sup>1</sup>, Richard Pearson<sup>1</sup>, Amy Strange<sup>1</sup>, Zhan Su<sup>1</sup>, Damjan Vukcevic<sup>1</sup>, Peter Donnelly<sup>1,2</sup>

DNA, Genotyping, Data QC and Informatics Group: Cordelia Langford<sup>3</sup>, Sarah E. Hunt<sup>3</sup>, Sarah Edkins<sup>3</sup>, Rhian Gwilliam<sup>3</sup>, Hannah Blackburn<sup>3</sup>, Suzannah J. Bumpstead<sup>3</sup>, Serge Dronov<sup>3</sup>, Matthew

Gillman<sup>3</sup>, Emma Gray<sup>3</sup>, Naomi Hammond<sup>3</sup>, Alagurevathi Jayakumar<sup>3</sup>, Owen T. McCann<sup>3</sup>, Jennifer Liddle<sup>3</sup>, Simon C. Potter<sup>3</sup>, Radhi Ravindrarah<sup>3</sup>, Michelle Ricketts<sup>3</sup>, Avazeh Tashakkori-Ghanbaria<sup>3</sup>, Matthew Waller<sup>3</sup>, Paul Weston<sup>3</sup>, Sara Widaa<sup>3</sup>, Pamela Whittaker<sup>3</sup>, Ines Barroso<sup>3</sup>, Panos Deloukas<sup>3</sup>

Publications Committee: Christopher G. Mathew (Chair)<sup>1,3</sup>, Jenefer M. Blackwell<sup>4,5</sup>, Matthew A. Brown<sup>7</sup>, Aiden Corvin<sup>9</sup>, Mark I. McCarthy<sup>19</sup>, Chris C.A. Spencer<sup>1</sup>

### WTCCC2 Affiliations:

<sup>1</sup>Wellcome Trust Centre for Human Genetics, Oxford, UK; <sup>2</sup>Department of Statistics, University of Oxford, Oxford, UK; <sup>3</sup>Wellcome Trust Sanger Institute, Wellcome Trust Genome Campus, Cambridge, UK; <sup>4</sup>Telethon Institute for Child Health Research, Centre for Child Health Research, University of Western Australia, Subiaco, Western Australia; <sup>5</sup>Cambridge Institute for Medical Research, University of Cambridge School of Clinical Medicine, Cambridge, UK; <sup>6</sup>University College London, UCL Mental Health Sciences Unit and UCL Institute of Cognitive Neurosciences, London, UK; <sup>7</sup>Diamantina Institute of Cancer, Immunology and Metabolic Medicine, Princess Alexandra Hospital, University of Queensland, Brisbane, Queensland, Australia; <sup>8</sup>Department of Epidemiology and Population Health, London School of Hygiene and Tropical Medicine, London and Department of Epidemiology and Public Health, University College London, UK; <sup>9</sup>Neuropsychiatric Genetics Research Group, Institute of Molecular Medicine, Trinity College Dublin, Dublin, Ireland; <sup>10</sup>Molecular and Physiological Sciences, The Wellcome Trust, London, UK; <sup>11</sup>Centre for Digestive Diseases, Queen Mary University of London, London, UK and Digestive Diseases Centre, Leicester Royal Infirmary, Leicester, UK and Department of Clinical Pharmacology, Old Road Campus, University of Oxford, Oxford, UK; <sup>12</sup>Clinical Neurosciences, St George's University of London, London, UK; <sup>13</sup>King's College London Department of Medical and Molecular Genetics, School of Medicine, Guy's Hospital, London, UK; <sup>14</sup>Biomedical Research Centre, Ninewells Hospital and Medical School, Dundee, UK; <sup>15</sup>King's College London Social, Genetic and Developmental Psychiatry Centre, Institute of Psychiatry; <sup>16</sup>University of Cambridge Department of Clinical Neurosciences, Addenbrooke's Hospital, Cambridge, UK; <sup>17</sup>NIHR Biomedical Research Centre at Moorfields Eye Hospital NHS Foundation Trust and UCL Institute of Ophthalmology, London, UK; <sup>18</sup>Department of Molecular Neuroscience, Institute of Neurology, Queen Square, London, UK; <sup>19</sup>Oxford Centre for Diabetes, Endocrinology and Metabolism (ICDEM), Churchill Hospital, Oxford, UK.

## Appendix A. Supporting information

Supplementary data

## References

1. van Os J., Hanssen M., Bijl R.V., Ravelli A. Strauss (1969) revisited: A psychosis continuum in the general population? *Schizophr Res.* 2000;45:11–20. [PubMed: 10978868]
2. van Os J., Hanssen M., Bijl R.V., Vollebergh W. Prevalence of psychotic disorder and community level of psychotic symptoms: An urban-rural comparison. *Arch Gen Psychiatry.* 2001;58:663–668. [PubMed: 11448373]
3. Jablensky A., Sartorius N., Ernberg G., Anker M., Korten A., Cooper J.E. Schizophrenia—manifestations, incidence and course in different cultures—a World Health Organization ten-country study. *Psychological Med.* 1992;20:1–97.
4. Murray R., Jones P., Susser E., Van Os J., Cannon M. *The Epidemiology of Schizophrenia.* Cambridge University Press; Cambridge: 2003.
5. Perälä J., Suvisaari J., Saarni S.I., Kuoppasalmi K., Isometsä E., Pirkola S. Lifetime prevalence of psychotic and bipolar I disorders in a general population. *Arch Gen Psychiatry.* 2007;64:19–28. [PubMed: 17199051]
6. Bogren M., Mattisson C., Isberg P.E., Nettelbladt P. How common are psychotic and bipolar

disorders? A 50-year follow-up of the Lundby population. *Nord J Psychiatry*. 2009;63:336–346. [PubMed: 19492244]

7. WHO The World Health Report 2001—Mental health, new understanding, new hope. *Eur J Psychiatry*. 2002;16 117–117.

8. Sullivan P.F., Daly M.J., O'Donovan M. Genetic architectures of psychiatric disorders: the emerging picture and its implications. *Nat Rev Genet*. 2012;13:537–551. [PubMed: 22777127]

9. Rees E., Kirov G., O'Donovan M.C., Owen M.J. De novo mutation in schizophrenia. *Schizophr Bull*. 2012;38:377–381. [PubMed: 22451492]

10. Sebat J., Levy D.L., McCarthy S.E. Rare structural variants in schizophrenia: One disorder, multiple mutations; one mutation, multiple disorders. *Trends Genet*. 2009;25:528–535. [PubMed: 19883952]

11. Malhotra D., McCarthy S., Michaelson J.J., Vacic V., Burdick K.E., Yoon S. High frequencies of de novo CNVs in bipolar disorder and schizophrenia. *Neuron*. 2011;72:951–963. [PubMed: 22196331]

12. Ripke S., Sanders A.R., Kendler K.S., Levinson D.F., Sklar P., Holmans P.A. Genome-wide association study identifies five new schizophrenia loci. *Nat Genet*. 2011;43:969–976. [PubMed: 21926974]

13. Alkelai A., Lupoli S., Greenbaum L., Giegling I., Kohn Y., Sarner-Kanyas K. Identification of new schizophrenia susceptibility loci in an ethnically homogeneous, family-based, Arab-Israeli sample. *FASEB J*. 2011;25:4011–4023. [PubMed: 21795503]

14. Athanasiu L., Mattingdal M., Kähler A.K., Brown A., Gustafsson O., Agartz I. Gene variants associated with schizophrenia in a Norwegian genome-wide study are replicated in a large European cohort. *J Psychiatr Research*. 2010;44:748–753. [PubMed: 20185149]

15. Green E.K., Grozeva D., Jones I., Jones L., Kirov G., Caesar S. The bipolar disorder risk allele at CACNA1C also confers risk of recurrent major depression and of schizophrenia. *Mol Psychiatry*. 2010;15:1016–1022. [PubMed: 19621016]

16. Kirov G., Zaharieva I., Georgieva L., Moskvina V., Nikolov I., Cichon S. A genome-wide association study in 574 schizophrenia trios using DNA pooling. *Mol Psychiatry*. 2008;14:796–803. [PubMed: 18332876]

17. Need A.C., Ge D., Weale M.E., Maia J., Feng S., Heinzen E.L. A genome-wide investigation of SNPs and CNVs in schizophrenia. *PLoS Genet*. 2009;5:e1000373. [PubMed: 19197363]

18. O'Donovan M.C., Craddock N., Norton N., Williams H., Peirce T., Moskvina V. Identification of loci associated with schizophrenia by genome-wide association and follow-up. *Nat Genet*. 2008;40:1053–1055. [PubMed: 18677311]

19. O'Donovan M.C., Norton N., Williams H., Peirce T., Moskvina V., Nikolov I. Analysis of 10 independent samples provides evidence for association between schizophrenia and a SNP flanking fibroblast growth factor receptor 2. *Mol Psychiatry*. 2009;14:30–36. [PubMed: 18813210]

20. Sachdev P., Smith J.S., Cathcart S. Schizophrenia-like psychosis following traumatic brain injury: A chart-based descriptive and case-control study. *Psychol Med*. 2001;31:231–239. [PubMed: 11232911]

21. Shi Y., Li Z., Xu Q., Wang T., Li T., Shen J. Common variants on 8p12 and 1q24.2 confer risk of schizophrenia. *Nat Genet*. 2011;43:1224–1227. [PubMed: 22037555]

22. Shifman S., Johannesson M., Bronstein M., Chen S.X., Collier D.A., Craddock N.J. Genome-wide association identifies a common variant in the reelin gene that increases the risk of schizophrenia only in women. *PLoS Genet*. 2008;4:e28. [PubMed: 18282107]

23. Shi J., Levinson D.F., Duan J., Sanders A.R., Zheng Y., Pe'er I. Common variants on chromosome 6p22.1 are associated with schizophrenia. *Nature*. 2009;460:753–757. [PubMed: 19571809]

24. Stefansson H., Ophoff R.A., Steinberg S., Andreassen O.A., Cichon S., Rujescu D. Common variants conferring risk of schizophrenia. *Nature*. 2009;460:744–747. [PubMed: 19571808]



25. Steinberg S., Mors O., Borglum A.D., Gustafsson O., Werge T., Mortensen P.B. Expanding the range of ZNF804A variants conferring risk of psychosis. *Mol Psychiatry*. 2011;16:59–66. [PubMed: 20048749]
26. Purcell S.M., Wray N.R., Stone J.L., Visscher P.M., O'Donovan M.C., Sullivan P.F. Common polygenic variation contributes to risk of schizophrenia and bipolar disorder. *Nature*. 2009;460:748–752. [PubMed: 19571811]
27. Irish Schizophrenia Genomics Consortium, Wellcome Trust Case-Control Consortium 2 Genome-wide association study implicates HLA-C\*01:02 as a risk factor at the major histocompatibility complex locus in schizophrenia. *Biol Psychiatry*. 2012;72:620–628. [PubMed: 22883433]
28. Sklar P., Ripke S., Scott L.J., Andreassen O.A., Cichon S., Craddock N. Large-scale genome-wide association analysis of bipolar disorder identifies a new susceptibility locus near ODZ4. *Nat Genet*. 2011;43:977–983. [PubMed: 21926972]
29. Chen D.T., Jiang X., Akula N., Shugart Y.Y., Wendland J.R., Steele C.J.M. Genome-wide association study meta-analysis of European and Asian-ancestry samples identifies three novel loci associated with bipolar disorder. *Mol Psychiatry*. 2011;18:195–205. [PubMed: 22182935]
30. Cichon S., Muhleisen T.W., Degenhardt F.A., Mattheisen M., Miro X., Strohmaier J. Genome-wide association study identifies genetic variation in neurocan as a susceptibility factor for bipolar disorder. *Am J Hum Genet*. 2011;88:372–381. [PubMed: 21353194]
31. Craddock N., Jones L., Jones I.R., Kirov G., Green E.K., Grozeva D., *et al.* (2008): Strong genetic evidence for a selective influence of GABA(A) receptors on a component of the bipolar disorder phenotype. *Mol Psychiatry* 15:146–153.
32. Djurovic S., Gustafsson O., Mattingsdal M., Athanasiu L., Bjella T., Tesli M. A genome-wide association study of bipolar disorder in Norwegian individuals, followed by replication in Icelandic sample. *J Affect Disord*. 2010;126:312–316. [PubMed: 20451256]
33. Ferreira M.A., O'Donovan M.C., Meng Y.A., Jones I.R., Ruderfer D.M., Jones L. Collaborative genome-wide association analysis supports a role for ANK3 and CACNA1C in bipolar disorder. *Nat Genet*. 2008;40:1056–1058. [PubMed: 18711365]
34. Jiang Y., Zhang H. Propensity score-based nonparametric test revealing genetic variants underlying bipolar disorder. *Genet Epidemiol*. 2011;35:125–132. [PubMed: 21254220]
35. Kerner B., Lambert C.G., Muthén B.O. Genome-wide association study in bipolar patients stratified by co-morbidity. *PLoS One*. 2011;6:e28477. [PubMed: 22205951]
36. Lee M.T.M., Chen C.H., Lee C.S., Chen C.C., Chong M.Y., Ouyang W.C. Genome-wide association study of bipolar I disorder in the Han Chinese population. *Mol Psychiatry*. 2011;16:548–556. [PubMed: 20386566]
37. Liu Y., Blackwood D.H., Caesar S., de Geus E.J., Farmer A., Ferreira M.A. Meta-analysis of genome-wide association data of bipolar disorder and major depressive disorder. *Mol Psychiatry*. 2011;16:2–4. [PubMed: 20351715]
38. Scott L.J., Muglia P., Kong X.Q., Guan W., Flickinger M., Upmanyu R. Genome-wide association and meta-analysis of bipolar disorder in individuals of European ancestry. *Proc Natl Acad Sci U S A*. 2009;106:7501–7506. [PubMed: 19416921]
39. Sklar P., Smoller J.W., Fan J., Ferreira M.A., Perlis R.H., Chambert K. Whole-genome association study of bipolar disorder. *Mol Psychiatry*. 2008;13:558–569. [PubMed: 18317468]
40. Smith E.N., Bloss C.S., Badner J.A., Barrett T., Belmonte P.L., Berrettini W. Genome-wide association study of bipolar disorder in European American and African American individuals. *Mol Psychiatry*. 2009;14:755–763. [PubMed: 19488044]
41. Smith E.N., Koller D.L., Panganiban C., Szelinger S., Zhang P., Badner J.A. Genome-wide association of bipolar disorder suggests an enrichment of replicable associations in regions near genes.

PLoS Genet. 2011;7:e1002134. [PubMed: 21738484]

42. Purcell S., Neale B., Todd-Brown K., Thomas L., Ferreira M., Bender D. PLINK: A toolset for whole-genome association and population-based linkage analysis. *Am J Hum Genet.* 2007;81:559–575. [PubMed: 17701901]
43. Ruderfer D.M., Kirov G., Chambert K., Moran J.L., Owen M.J., O'Donovan M.C. A family-based study of common polygenic variation and risk of schizophrenia. *Mol Psychiatry.* 2011;16:887–888. [PubMed: 21483432]
44. Lander E.S. Initial impact of the sequencing of the human genome. *Nature.* 2011;470:187–197. [PubMed: 21307931]
45. Bouzigon E., Forabosco P., Koppelman G.H., Cookson W.O.C.M., Dizier M.H., Duffy D.L. Meta-analysis of 20 genome-wide linkage studies evidenced new regions linked to asthma and atopy. *Eur J Hum Genet.* 2010;18:700–706. [PubMed: 20068594]
46. Lewis C.M., Ng M.Y., Butler A.W., Cohen-Woods S., Uher R., Pirlo K. Genome-wide association study of major recurrent depression in the U.K. population. *Am J Psychiatry.* 2010;167:949–957. [PubMed: 20516156]
47. Linscott R.J., van Os J. Systematic reviews of categorical versus continuum models in psychosis: Evidence for discontinuous subpopulations underlying a psychometric continuum. Implications for DSM-V, DSM-VI, and DSM-VII. *Annu Rev Clin Psychol.* 2010;6:391–419. [PubMed: 20192792]
48. Kempton M.J., Geddes J.R., Ettinger U., Williams S.C., Grasby P.M. Meta-analysis, database, and meta-regression of 98 structural imaging studies in bipolar disorder. *Arch Gen Psychiatry.* 2008;65:1017–1032. [PubMed: 18762588]
49. Bramon E., Sham P. The common genetic liability between schizophrenia and bipolar disorder: A review. *Curr Psychiatry Rep.* 2001;3:332–337. [PubMed: 11470041]
50. Bramon E., Sham P. The shared genetic architecture which underlies schizophrenia and bipolar disorder. In: McDonald C., Schulze K., Murray R., Wright P., editors. *Schizophrenia, Challenging the Orthodox.* Taylor & Francis; London: 2004. pp. 173–181.
51. Sullivan P., Cichon S., Craddock N., Daly M., Faraone S.V., Gejman P.V. A framework for interpreting genome-wide association studies of psychiatric disorders. *Mol Psychiatry.* 2009;14:10–17. [PubMed: 19002139]
52. Williams N.M., Zaharieva I., Martin A., Langley K., Mantripragada K., Fossdal R. Rare chromosomal deletions and duplications in attention-deficit hyperactivity disorder: A genome-wide analysis. *Lancet.* 2010;376:1401–1408. [PubMed: 20888040]
53. Owen M.J., O'Donovan M.C., Thapar A., Craddock N. Neurodevelopmental hypothesis of schizophrenia. *Br J Psychiatry.* 2011;198:173–175. [PubMed: 21357874]
54. Huang J., Perlis R.H., Lee P.H., Rush A.J., Fava M., Sachs G.S. Cross-disorder genomewide analysis of schizophrenia, bipolar disorder, and depression. *Am J Psychiatry.* 2010;167:1254–1263. [PubMed: 20713499]
55. Williams H.J., Craddock N., Russo G., Hamshere M.L., Moskvina V., Dwyer S. Most genome-wide significant susceptibility loci for schizophrenia and bipolar disorder reported to date cross-traditional diagnostic boundaries. *Hum Mol Genet.* 2011;20:387–391. [PubMed: 21037240]
56. Sullivan P.F. The Psychiatric GWAS Consortium: Big science comes to psychiatry. *Neuron.* 2010;68:182–186. [PubMed: 20955924]
57. Wray N.R., Pergadia M.L., Blackwood D.H.R., Penninx BWJH, Gordon S.D., Nyholt D.R. Genome-wide association study of major depressive disorder: New results, meta-analysis, and lessons learned. *Mol Psychiatry.* 2012;17:36–48. [PubMed: 21042317]
58. Owen M.J. Intellectual disability and major psychiatric disorders: A continuum of neurodevelopmental causality. *Br J Psychiatry.* 2012;200:268–269. [PubMed: 22474230]

59. Sullivan P.F. Puzzling over schizophrenia: schizophrenia as a pathway disease. *Nat Med*. 2012;18:210–211. [PubMed: 22310687]
60. American Psychiatric Association . American Psychiatric Association; Washington, DC: 2000. Diagnostic and Statistical Manual of Mental Disorders, 4th ed., text rev.
61. Wing J.K., Babor T., Brugha T., Burke J., Cooper J.E., Giel R. SCAN. Schedules for Clinical Assessment in Neuropsychiatry. *Arch Gen Psychiatry*. 1990;47:589–593. [PubMed: 2190539]
62. Endicott J., Spitzer R.L. A diagnostic interview: The schedule for affective disorders and schizophrenia. *Arch Gen Psychiatry*. 1978;35:837–844. [PubMed: 678037]
63. Spitzer R.L., Williams J.B.W., Gibbon M., First M.B. The Structured Clinical Interview for DSM-III-R (SCID). 1. History, rationale, and description. *Arch Gen Psychiatry*. 1992;49:624–629. [PubMed: 1637252]
64. Marchini J., Howie B., Myers S., McVean G., Donnelly P. A new multipoint method for genome-wide association studies by imputation of genotypes. *Nat Genet*. 2007;39:906–913. [PubMed: 17572673]
65. Wellcome Trust Case Control Consortium Genome-wide association study of 14,000 cases of seven common diseases and 3,000 shared controls. *Nature*. 2007;447:661–U667. [PubMed: 17554300]
66. Wigginton J.E., Abecasis G.R. PEDSTATS: Descriptive statistics, graphics and quality assessment for gene mapping data. *Bioinformatics*. 2005;21:3445–3447. [PubMed: 15947021]
67. Morris J.A., Randall J.C., Maller J.B., Barrett J.C. Evoker: A visualization tool for genotype intensity data. *Bioinformatics*. 2010;26:1786–1787. [PubMed: 20507892]
68. Patterson N., Price A.L., Reich D. Population structure and eigenanalysis. *PLoS Genet*. 2006;2:e190. [PubMed: 17194218]
69. Price A.L., Patterson N.J., Plenge R.M., Weinblatt M.E., Shadick N.A., Reich D. Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet*. 2006;38:904–909. [PubMed: 16862161]
70. Devlin B., Roeder K., Wasserman L. Genomic control, a new approach to genetic-based association studies. *Theor Popul Biol*. 2001;60:155–166. [PubMed: 11855950]
71. Dudbridge F. Likelihood-based association analysis for nuclear families and unrelated subjects with missing genotype data. *Hum Hered*. 2008;66:87–98. [PubMed: 18382088]
72. Hindorff L.A., Sethupathy P., Junkins H.A., Ramos E.M., Mehta J.P., Collins F.S. Potential etiologic and functional implications of genome-wide association loci for human diseases and traits. *Proc Natl Acad Sci*. 2009;106:9362–9367. [PubMed: 19474294]
73. The 1000 Genomes Project Consortium A map of human genome variation from population-scale sequencing. *Nature*. 2010;467:1061–1073. [PubMed: 20981092]
74. Johnson A.D., Handsaker R.E., Pulit S.L., Nizzari M.M., O'Donnell C.J., de Bakker P.I. SNAP: A web-based tool for identification and annotation of proxy SNPs using HapMap. *Bioinformatics*. 2008;24:2938–2939. [PubMed: 18974171]
75. Steinberg S., de Jong S., Andreassen O.A., Werge T., Børghlum A.D., Mors O. Common variants at VRK2 and TCF4 conferring risk of schizophrenia. *Hum Mol Genet*. 2011;20:4076–4081. [PubMed: 21791550]
76. McMahon F.J., Akula N., Schulze T.G., Muglia P., Tozzi F., Detera-Wadleigh S.D. Meta-analysis of genome-wide association data identifies a risk locus for major mood disorders on 3p21.1. *Nat Genet*. 2010;42:128–131. [PubMed: 20081856]
77. Liou Y.-J., Wang H.-H., Lee M.-T.M., Wang S.-C., Chiang H.-L., Chen C.-C. Genome-wide association study of treatment refractory schizophrenia in Han Chinese. *PLoS One*. 2012;7:e33598. [PubMed: 22479419]

78. Stefansson H., Rujescu D., Cichon S., Pietilainen O.P., Ingason A., Steinberg S. Large recurrent microdeletions associated with schizophrenia. *Nature*. 2008;455:232–236. [PubMed: 18668039]
79. Li T., Li Z., Chen P., Zhao Q., Wang T., Huang K. Common variants in major histocompatibility complex region and TCF4 gene are significantly associated with schizophrenia in Han Chinese. *Biol Psychiatry*. 2010;68:671–673. [PubMed: 20673877]
80. Yue W.-H., Wang H.-F., Sun L.-D., Tang F.-L., Liu Z.-H., Zhang H.-X. Genome-wide association study identifies a susceptibility locus for schizophrenia in Han Chinese at 11p11.2. *Nat Genet*. 2011;43:1228–1231. [PubMed: 22037552]
81. Wang K.-S., Liu X.-F., Aragam N. A genome-wide meta-analysis identifies novel loci associated with schizophrenia and bipolar disorder. *Schizophr Res*. 2010;124:192–199. [PubMed: 20889312]
82. Bergen S.E., O'Dushlaine C.T., Ripke S., Lee P.H., Ruderfer D.M., Akterin S. Genome-wide association study in a Swedish population yields support for greater CNV and MHC involvement in schizophrenia compared with bipolar disorder. *Mol Psychiatry*. 2012;17:880–886. [PubMed: 22688191]
83. Schulze T.G., Detera-Wadleigh S.D., Akula N., Gupta A., Kassem L., Steele J. Two variants in Ankyrin 3 (ANK3) are independent genetic risk factors for bipolar disorder. *Mol Psychiatry*. 2009;14:487–491. [PubMed: 19088739]
84. Lett T.A.P., Zai C.C., Tiwari A.K., Shaikh S.A., Likhodi O., Kennedy J.L. ANK3, CACNA1C and ZNF804A gene variants in bipolar disorders and psychosis subphenotype. *World J Biol Psychiatry*. 2011;12:392–397. [PubMed: 21767209]
85. Takata A., Kim S.H., Ozaki N., Iwata N., Kunugi H., Inada T. Association of ANK3 with bipolar disorder confirmed in East Asia. *Am J Med Genet B Neuropsychiatr Genet*. 2011;156:312–315. [PubMed: 21438140]
86. Tesli M., Koefoed P., Athanasiu L., Mattingsdal M., Gustafsson O., Agartz I. Association analysis of ANK3 gene variants in Nordic bipolar disorder and schizophrenia case–control samples. *Am J Med Genet B Neuropsychiatr Genet*. 2011;156:969–974. [PubMed: 21972176]
87. Guan F., Wei S., Feng J., Zhang C., Xing B., Zhang H. Association study of a new schizophrenia susceptibility locus of 10q24.32–33 in a Han Chinese population. *Schizophr Res*. 2012;138:63–68. [PubMed: 22520855]
88. Smith R.L., Knight D., Williams H., Dwyer S., Richards A., Kirov G. Analysis of neurogranin (NRGN) in schizophrenia. *Am J Med Genet B Neuropsychiatr Genet*. 2011;156:532–535. [PubMed: 21538840]
89. Nyegaard M., Demontis D., Foldager L., Hedemand A., Flint T.J., Sorensen K.M. CACNA1C (rs1006737) is associated with schizophrenia. *Mol Psychiatry*. 2010;15:119–121. [PubMed: 20098439]
90. Baum A.E., Akula N., Cabanero M., Cardona I., Corona W., Klemens B. A genome-wide association study implicates diacylglycerol kinase eta (DGKH) and several other genes in the etiology of bipolar disorder. *Mol Psychiatry*. 2007;13:197–207. [PubMed: 17486107]
91. Baum A.E., Hamshere M., Green E., Cichon S., Rietschel M., Nothen M.M. Meta-analysis of two genome-wide association studies of bipolar disorder reveals important points of agreement. *Mol Psychiatry*. 2008;13:466–467. [PubMed: 18421293]
92. Yosifova A., Mushiroda T., Kubo M., Takahashi A., Kamatani Y., Kamatani N. Genome-wide association study on bipolar disorder in the Bulgarian population. *Genes Brain Behav*. 2011;10:789–797. [PubMed: 21771265]
93. Ollila H.M., Soronen P., Silander K., Palo O.M., Kieseppa T., Kaunisto M.A. Findings from bipolar disorder genome-wide association studies replicate in a Finnish bipolar family-cohort. *Mol Psychiatry*. 2008;14:351–353. [PubMed: 19308021]
94. Tesli M., Athanasiu L., Mattingsdal M., Kähler A.K., Gustafsson O., Andreassen B.K. Association analysis of PALB2 and BRCA2 in bipolar disorder and schizophrenia in a Scandinavian case–control

sample. *Am J Med Genet B Neuropsychiatr Genet.* 2010;153B:1276–1282. [PubMed: 20872766]

95. Mühleisen T.W., Mattheisen M., Strohmaier J., Degenhardt F., Priebe L., Schultz C.C. Association between schizophrenia and common variation in neurocan (NCAN), a genetic risk factor for bipolar disorder. *Schizophr Res.* 2012;138:69–73. [PubMed: 22497794]

96. Cardno A., Gottesman I. Twin studies of schizophrenia: From bow-and-arrow concordances to star wars mx and functional genomics. *Am J Med Genet B Neuropsychiatr Genet.* 2000;97:12–17.

97. Cardno A.G., Rijdsdijk F.V., Sham P.C., Murray R.M., McGuffin P. A twin study of genetic relationships between psychotic symptoms. *Am J Psychiatry.* 2002;159:539–545. [PubMed: 11925290]

98. Sullivan P.F., Kendler K.S., Neale M.C. Schizophrenia as a complex trait—evidence from a meta-analysis of twin studies. *Arch Gen Psychiatry.* 2003;60:1187–1192. [PubMed: 14662550]

99. McGuffin P. Gene polymorphisms and behavior. *Pediatr Blood Cancer.* 2007;48:736–737. [PubMed: 17311348]

100. Lichtenstein P., Yip B.H., Bjork C., Pawitan Y., Cannon T.D., Sullivan P.F. Common genetic determinants of schizophrenia and bipolar disorder in Swedish families: A population-based study. *Lancet.* 2009;373:234–239. [PubMed: 19150704]

101. Antonarakis S.E., Chakravarti A., Cohen J.C., Hardy J. Mendelian disorders and multifactorial traits: The big divide or one for all? *Nat Rev Genet.* 2010;11:380–384. [PubMed: 20395971]

102. Hardy J., Low N., Singleton A. Whole genome association studies: deciding when persistence becomes perseverance. *Am J Med Genet B Neuropsychiatr Genet.* 2008;147B:131–133. [PubMed: 17541974]

103. Hardy J., Low N.C. Genes and environment in psychiatry: Winner's curse or cure? *Arch Gen Psychiatry.* 2011;68:455–456. [PubMed: 21536973]

104. Owen M.J., Craddock N., O'Donovan M.C. Suggestion of roles for both common and rare risk variants in genome-wide studies of schizophrenia. *Arch Gen Psychiatry.* 2010;67:667–673. [PubMed: 20603448]

105. Bellenguez C., Bevan S., Gschwendtner A., Spencer C.C., Burgess A.I., Pirinen M. Genome-wide association study identifies a variant in HDAC9 associated with large vessel ischemic stroke. *Nat Genet.* 2012;44:328–333. [PubMed: 22306652]

106. Sawcer S., Hellenthal G., Pirinen M., Spencer C.C., Patsopoulos N.A., Moutsianas L. Genetic risk and a primary role for cell-mediated immune mechanisms in multiple sclerosis. *Nature.* 2011;476:214–219. [PubMed: 21833088]

107. Spencer C.C., Plagnol V., Strange A., Gardner M., Paisan-Ruiz C., Band G. Dissection of the genetics of Parkinson's disease identifies an additional association 5' of SNCA and multiple associated haplotypes at 17q21. *Hum Mol Genet.* 2011;20:345–353. [PubMed: 21044948]

108. Burton P.R., Hansell A.L., Fortier I., Manolio T.A., Khoury M.J., Little J. Size matters: Just how big is BIG?: Quantifying realistic sample size requirements for human genome epidemiology. *Int J Epidemiol.* 2009;38:263–273. [PubMed: 18676414]

109. Cichon S., Craddock N., Daly M., Faraone S.V., Gejman P.V., Kelsoe J. Genomewide association studies: History, rationale, and prospects for psychiatric disorders. *Am J Psychiatry.* 2009;166:540–556. [PubMed: 19339359]

110. Kirov G., O'Donovan M.C., Owen M.J. Finding schizophrenia genes. *J Clin Invest.* 2005;115:1440–1448. [PubMed: 15931379]

111. Kirov G., Rujescu D., Ingason A., Collier D.A., O'Donovan M.C., Owen M.J. Neurexin 1 (NRXN1) deletions in schizophrenia. *Schizophr Bull.* 2009;35:851–854. [PubMed: 19675094]

112. Vacic V., McCarthy S., Malhotra D., Murray F., Chou H.H., Peoples A. Duplications of the neuropeptide receptor gene VIPR2 confer significant risk for schizophrenia. *Nature.* 2011;471:499–503.



[PubMed: 21346763]

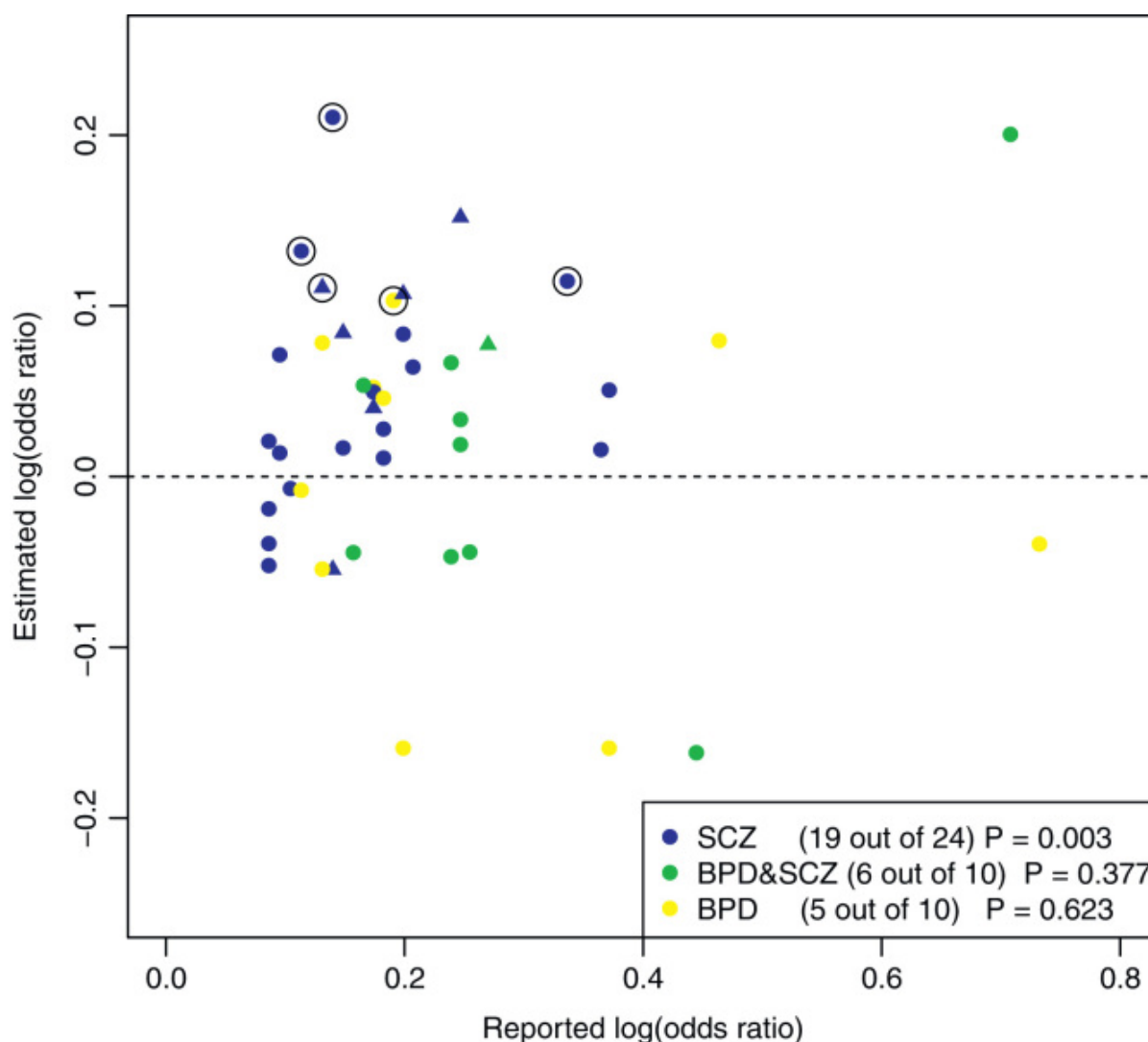
113. McCarthy S.E., Makarov V., Kirov G., Addington A.M., McClellan J., Yoon S. Microduplications of 16p11.2 are associated with schizophrenia. *Nat Genet.* 2009;41:1223–1227. [PubMed: 19855392]

114. Ioannidis J.P. Why most discovered true associations are inflated. *Epidemiology.* 2008;19:640–648. [PubMed: 18633328]

115. Zollner S., Pritchard J.K. Overcoming the winner's curse: Estimating penetrance parameters from case-control data. *Am J Hum Genet.* 2007;80:605–615. [PubMed: 17357068]

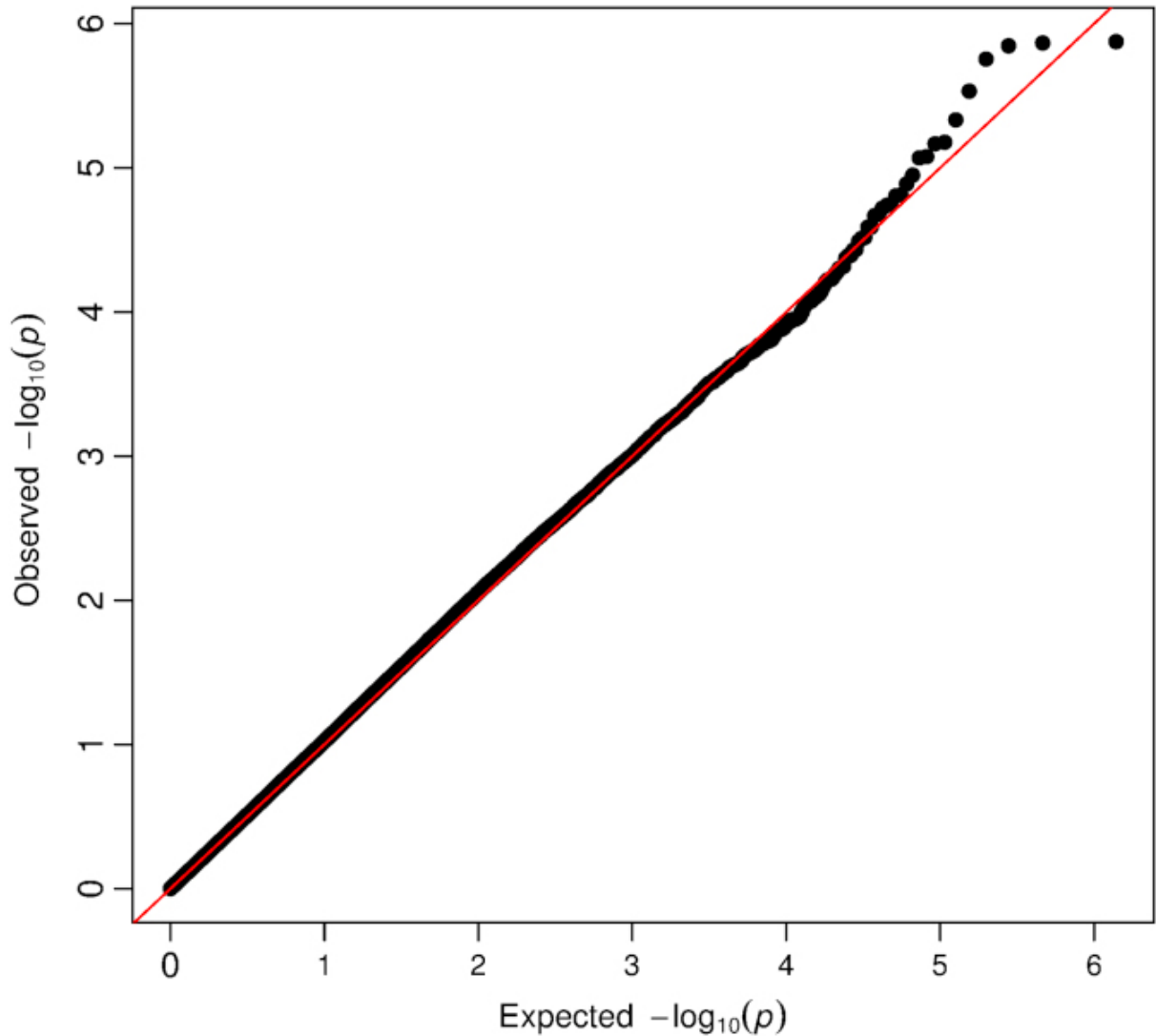
## Figures and Tables

**Figure 1**



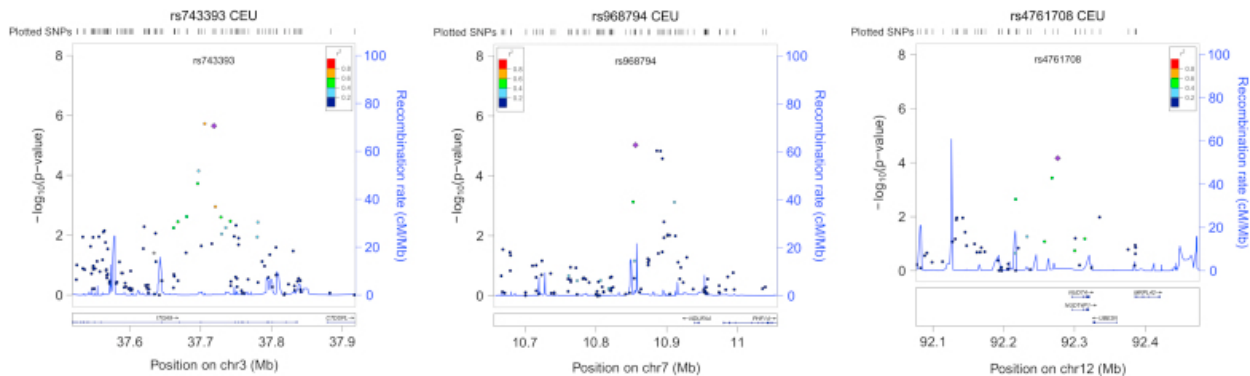
Replication of previously published associated schizophrenia (SCZ) and bipolar disorder (BPD) loci. Reported odds ratios from the literature (x axis) plotted against the odds ratios estimated from our data as listed in [Table 1](#) (y axis). The dotted line indicates an odds ratio of 1 in our data. Points above the line indicate the same direction of effect in previous studies and our data. Black circles indicate single nucleotide polymorphisms that replicate ( $p$  one-sided  $< .05$ ) in our study. Triangles denote single nucleotide polymorphisms within the major histocompatibility complex region. Sign tests for an enrichment of effect in the same direction are presented for loci previously associated with schizophrenia, bipolar disorder, or both. Additional details on the sign tests are available in [Table S4](#) of [Supplement 1](#).

**Figure 2**



Quantile-quantile plots of the distribution of observed and expected  $p$  value at all autosomal single nucleotide polymorphisms passing quality control. The genomic control lambda value was 1.03. This represents analyses conducted on our entire discovery sample ( $n = 4835$ ).

Figure 3



Regional association plots for the three single nucleotide polymorphisms (SNPs) reported in Table 2. The plots show the evidence of association in our discovery data at genotyped SNPs. The SNP listed in Table 2 is indicated with a diamond, and flanking SNPs (circles) are colored according to their correlations ( $r^2$ ) with this SNP measured in HapMap. CEU, Utah residents (CEPH) with Northern and Western European ancestry.

Table 1

Evidence in Our Discovery Data at Loci Previously Reported to Be Associated with SCZ and/or BPD

Chr	Position	Genes	SNP Risk	OR Phen	Previous	Best Tag	$r^2$ 1000	Freq	$p$
-----	----------	-------	----------	---------	----------	----------	------------	------	-----

	(b36)		Allele			Studies		Genomes		
1	98502934	<i>MIR137</i>	rs1625579-T	1.12	SCZ	<a href="#">(12)</a>	rs1782810	.975	.8034	.04
1	167903079	<i>BRP44, DCAF6</i>	rs10489202-A	1.23	SCZ	<a href="#">(21)</a>	rs10489202	1.000	.2312	.3
1	243608967	—	rs6703335-G	1.09	SCZ	<a href="#">(12)</a>	rs6703335	1.000	.4324	.4
2	58222928	<i>VRK2</i>	rs2312147-C	1.09	SCZ	<a href="#">(24,75)</a>	rs1518395	1.000	.6172	.7
2	193984621	<i>PCGEM1</i>	rs17662626-A	1.2	SCZ	<a href="#">(12)</a>	rs17666314	.633	.942	.8
3	36862980	—	rs4624519-T	1.09	SCZ	<a href="#">(12)</a>	rs4624519	1.000	.6264	.7
3	52821011	<i>ITIH1, NEK4</i>	rs1042779-A	1.19	BPD	<a href="#">(38,76)</a>	rs1042779	1.000	.6237	.3
3	52855229	<i>ITIH3, ITIH4</i>	rs2239547-T	1.1	SCZ	<a href="#">(12)</a>	rs1573815	.786	.7678	.2
3	180550702	<i>FXR1</i>	rs6782299-T	1.1	SCZ	<a href="#">(26)</a>	rs6782299	1.000	.7294	.8
4	103457418	<i>NFKB1</i>	rs230529-T	1.45	SCZ	<a href="#">(77)</a>	rs230529	1.000	.3895	.3
5	94154588	<i>MCTP1</i>	rs17418283-C	1.21	BPD	<a href="#">(38)</a>	rs255339	.834	.2968	.04
6	27143883	<i>MHC</i>	rs13194053-T	1.22	SCZ	<a href="#">(23,26)</a>	rs7745603	.866	.7697	.1
6	27248931	<i>MHC</i>	rs6932590-T	1.16	SCZ	<a href="#">(24,26,78,79)</a>	rs6938200	.874	.7947	.3
6	28227604	<i>MHC,NKAPL</i>	rs1635-G	1.28	SCZ	<a href="#">(80)</a>	rs2299030	.909	.03693	.3
6	30174131	<i>MHC,TRIM26</i>	rs2021722-C	1.15	SCZ	<a href="#">(12)</a>	rs2517611	1.000	.8058	.4
6	32172993	<i>MHC, NOTCH4</i>	rs3131296-G	1.19	SCZ	<a href="#">(24,78)</a>	rs2071278	1.000	.8534	.
6	32602269	<i>MHC, HLA-DQA1</i>	rs9272219-G	1.14	SCZ	<a href="#">(23)</a>	rs9272219	1.000	.7105	.07
6	89732101	<i>GABRR1</i>	rs12201676-	1.31	SCZ,	<a href="#">(81)</a>	rs12201676	1.000	.2452	.3

			C		BPD					
7	145959243	CNTNAP2	rs802568-C	2.03	SCZ, BPD	(81)	rs802568	1.000	.02746	.1
7	156048649	LOC393076	rs10949808-T	1.27	SCZ, BPD	(81)	rs10949808	1.000	.397	.2
8	4180844	CSMD1	rs10503253-A	1.16	SCZ	(12)	rs10103330	1.000	.1993	.8
8	38031345	LSM1, WHSC1L1	rs16887244-A	1.19	SCZ	(21)	rs16887244	1.000	.7614	.2
8	58840924	FAM110B	rs1992045-T	1.56	SCZ, BPD	(81)	rs1992045	1.000	.07363	.1
8	89760311	MMP16	rs7004633-A	1.11	SCZ	(12,82)	rs7005110	1.000	.1925	.9
9	121359286	ASTN2	rs11789399-G	1.29	SCZ, BPD	(81)	rs11789399	1.000	.4962	.2
10	62179812	ANK3	rs10994336-T	1.45	BPD	(33,38,83-86)	rs3808943	.858	.05389	.2
10	62279124	ANK3	rs10994397-T	1.22	BPD	(28)	rs3808943	.858	.05389	.2
10	104775908	CNNM2	rs7914558-G	1.22	SCZ	(12,87)	rs10509757	1.000	.5927	.1
10	104906211	NT5C2	rs11191580-T	1.2	SCZ	(12,87)	rs11191580	1.000	.9169	.9
11	17160148	—	rs4356203-G	1.09	SCZ	(12)	rs4356203	1.000	.4236	.3
11	66551002	Multiple	rs10896135-G	1.12	BPD	(28)	rs4930390	1.000	.2557	.9
11	79077193	ODZ4	rs12576775-G	1.14	BPD	(28)	rs7932890	1.000	.1737	.2
11	98125404	CNTN5	rs2509843-C	1.27	SCZ, BPD	(81)	rs2848547	1.000	.3711	.2
11	124606285	NRGN	rs12807809-T	1.15	SCZ	(24,80,89)	rs1939214	.972	.8282	.06
12	2345295	CACNA1C	rs1006737-A	1.18	SCZ, BPD	(15,33,37,38,90)	rs1006737	1.000	.3265	.3

12	2419896	<i>CANCNA1C</i>	rs4765913-A	1.14	BPD	<a href="#">(28)</a>	rs4765914	.935	.8061	.4
13	42653437	<i>DGKH</i>	rs1012053-A	1.59	BPD	<a href="#">(91–93)</a>	rs1170188	.806	.8133	.
13	106651661	<i>DAOA</i>	rs4996815-C	1.28	SCZ, BPD	<a href="#">(82)</a>	rs4996815	1.000	.3576	.7
15	38995491	Multiple	rs12899449-A	1.2	BPD	<a href="#">(33)</a>	rs12899449	1.000	.7063	.4
16	23634026	Multiple	rs420259-A	2.08	BPD	<a href="#">(34,66,94,95)</a>	rs420259	1.000	.7306	.5
18	52752017	<i>CCDC68</i>	rs12966547-G	1.4	SCZ	<a href="#">(12)</a>	rs11874716	1.000	.567	.05
18	53058238	<i>TCF4</i>	rs17594526-T	1.44	SCZ	<a href="#">(26)</a>	rs17594526	1.000	.0321	.5
19	19361735	<i>NCAN</i>	rs1064395-A	1.17	SCZ, BPD	<a href="#">(30,96)</a>	rs1064395	1.000	.1639	.5
20	19852503	<i>RIN2</i>	rs6046396-G	1.28	SCZ, BPD	<a href="#">(82)</a>	rs6046396	1.000	.3015	.

Evidence for association at the best marker, as measured by correlation ( $r^2$ ) in the 1000 Genomes Project data [\(74\)](#), of previously implicated loci. Where the previously reported single nucleotide polymorphism was not typed in our data, we oriented the haplotypes to estimate the odds ratio for the risk allele allowing a one-sided comparison.

BPD, bipolar disorder; Chr, chromosome; CI, confidence interval; Freq, frequency; OR, odds ratio; Phen, phenotype; SCZ, schizophrenia.

<sup>a</sup>One-sided  $p < .05$  (for replication with the same risk allele).

**Table 2**

Loci Where the Strongest Evidence of Association Was Found

Chr	SNP (Risk Allele)	Position (Build 36)	Discovery $p$ OR (95% CI)	Replication $p$ OR (95% CI)	Combined $p$ OR (95% CI)	log10BF Same	log10BF Related
3	rs743393-C	37718972	9.40E-06 1.301 (1.157–1.462)	.039 1.043 (1.002–1.085)	7.40E-04 1.067 (1.028–1.108)	1.43	2.64
7	rs968794-A	10856010	8.30E-07 1.324 (1.184–1.480)	.014 1.050 (1.010–1.091)	8.40E-05 1.076 (1.037–1.116)	2.3	3.87
12	rs4761708-C	92276217	1.40E-04 1.291 (1.130–1.476)	.024 1.058 (1.007–1.110)	6.40E-04 1.083 (1.034–1.133)	1.56	2.18



Evidence of association across the discovery sample and the Psychiatric GWAS Consortium and SGENE replication collections as also shown as a forest plot in [Figure S3](#) in [Supplement 1](#). We report  $p$  values, ORs (and 95% CIs) for these cohorts and the combined evidence (using an inverse variance weighted fixed effects approach). We also show the  $\log_{10}$ BF for models assuming effects to be either the same (fixed) or related (allowing heterogeneity). Note that the discovery sample ORs and  $p$  values presented in this table are based on a subset of our sample that did not overlap with the replication cohorts; therefore, the  $p$  values may vary from those presented in [Figure S2](#) in [Supplement 1](#), which includes our entire discovery data set.

BF, Bayes factor; Chr, chromosome; CI, confidence interval; OR, odds ratio; SNP, single nucleotide polymorphism.